BEFORE THE
FEDERAL COMMUNICATIONS COMMISSION
WASHINGTON, D.C. 20554

In the Matter of Closed Captioning of Internet Protocol-Delivered Video Programming: Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010	)	MB Docket No. 11-154

COMMENTS OF ROVI CORPORATION

Adam Powers
Vice President of Technology
Rovi Corporation
2830 De La Cruz Blvd
Santa Clara, CA 95050
apowers@rovicorp.com

October 18, 2011

TABLE OF CONTENTS

BEFORE THE
FEDERAL COMMUNICATIONS COMMISSION
WASHINGTON, D.C. 20554

In the Matter of Closed Captioning of Internet Protocol-Delivered Video Programming: Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010	)	MB Docket No. 11-154

COMMENTS OF ROVI CORPORATION

SUMMARY

The VPO may not be the most likely to know whether captions are required, and the owner of the video programming may not have a direct business relationship with the VPD/VPP, therefore the Commission should have a "safe harbor" for relying upon either certifications and/or third party data.

In order to minimize costs, burdens and implementation time, the Commission should adopt SMPTE-TT as the interchange format for captions, limited to CEA-608 and CEA-708 functionality. In the event the Commission concludes that specifying a single interchange format, it should specify a small set of interchange formats, to limit costs, burdens and implementation time. The Commission should not specify caption delivery formats for IP-delivered programming.

Apparatus covered by the regulation should be limited to devices that receive or play back programming which is transmitted. Furthermore, such apparatus should be limited to devices which are capable of receiving or playing back IP-delivered programming, and have such capability enabled, and such capability is not ancillary to the device's essential utility.

Finally, the Commission should "grandfather" certain programming which have been received and ingested by VPD/VPPs long before the effective date, but which were never supplied with caption data.

BEFORE THE
FEDERAL COMMUNICATIONS COMMISSION
WASHINGTON, D.C. 20554

In the Matter of Closed Captioning of Internet Protocol-Delivered Video Programming: Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010	)	MB Docket No. 11-154

COMMENTS OF ROVI CORPORATION

The Commission has sought comment on rules proposed to implement certain provisions of the Twenty-First Century Communications and Video Accessibility Act of 2010 ("CVAA") related to closed captioning of video programming delivered via the Internet.[1]

I. INTRODUCTION

Rovi Corporation ("Rovi"), founded in 1983 as Macrovision, is an industry leading provider of digital entertainment technologies for businesses in the consumer electronic, cable and satellite and entertainment markets across the world. The company is focused on developing entertainment technology that helps consumers sort through the numerous programming options available to find television shows and movies to watch. With its acquisitions of Gemstar-TV Guide and Sonic Solutions, Rovi is a preeminent provider of entertainment content distribution and navigation technologies, entertainment information and intellectual property.

Rovi Corporation is an industry-leading provider of both consumer-facing and professional products and services world-wide. Our businesses include services and technologies such as electronic program guide products, home and professional content authoring systems, Internet content delivery services, and Internet receiver solutions.

Rovi operates a "white label"[2] service (known as Rovi Entertainment Solutions, or "RES") an online video provider which licenses content and operates the back-end services, billing,

[1] *See Closed Captioning of Internet Protocol-Delivered Video Programming: Implementation of the Twenty-First Century Communications and Video Accessibility Act of 2010*, Notice of Proposed Rulemaking, MB Docket No. 11-154, FCC 11-138 (rel. Sep. 19, 2011) ("*NPRM*").

infrastructure and content delivery as a for a number of retail stores, including BestBuy, Blockbuster, Warner Brothers, Lionsgate and others. Additionally, Rovi's products include a widely-adopted video format (DivX), a metadata service for television, and a number of program guides and video players for televisions, tablets, smartphones and other devices. With these varied businesses, Rovi products and services are affected by content interchange mechanisms ("ingest") and captioning decoding/display technologies.

Because of its diverse business interests, Rovi participates in a variety of industry standards-setting organizations and consortia. Additionally, Rovi is proud to participate as a member of the VPAAC. With this experience as a participant in the process and stakeholder in the eventual regulations, Rovi respectfully submits these comments.

II. VIDEO PROGRAMMING OWNERS ARE OFTEN NOT THE MOST LIKELY TO KNOW WHETHER CAPTIONS ARE REQUIRED

Unfortunately, video programming owners ("VPOs") are not the most likely to know whether captions are required. The Commission "presume[s] that VPOs are in the best position to know if captions are required," and seeks comment on this presumption.[3] The Commission proposes to require that content be supplied either with captions or a certification indicating why captions are not required.[4] The Commission also seeks comment on how such certifications should be treated and supplied.[5]

A. LARGE, INTEGRATED VPOS OPERATE SEPARATE BROADCAST AND LICENSING BUSINESSES

Major studios operate several different, separate businesses. Content is produced, then generally licensed for broadcast by one business unit and licensed for video-on-demand (including IP) and home video by another business unit.[6]

---

[2] Rovi does not sell this service directly to consumers. *See* white-label product, http://en.wikipedia.org/wiki/White-label_product (last visited June 8, 2011).
[3] *NPRM* at 35.
[4] *Id.*
[5] *Id.* at 35-36.
[6] It should be noted that Rovi is not a studio, and we merely offer our perspective on studio business organization as an "outsider," and in general terms. We presume that the VPOs will describe to the Commission specific operational issues.

One business unit (broadcast licensing) licenses content for broadcast and in some cases, it may have knowledge of broadcast dates – but in other cases (especially library content) it has only vague knowledge of when content will be broadcast. An additional complication occurs when content is licensed on a nonexclusive basis, yielding many different broadcasters each with a time window during which the content may be broadcast.

Another business unit licenses content for on-demand distribution, which generally includes distribution via IP networks. In this case, a license is granted for a (renewable) period of time, and the distributor supplies the content to a consumer as a consumer requests (and/or pays for) it. The VPO has no knowledge of when (or if) such distributions take place until long after such a distribution occurred (*e.g.*, when remuneration is made by the video programming distributor ("VPD" to the VPO). Moreover, such licensing arrangements are typically nonexclusive, which introduces an additional layer of complexity to understanding when content is transmitted to consumers.

VPO's complicated businesses frequently distance IP-delivered content licensing at arm's length from broadcast-delivered content licensing, and there are significant structural obstacles to one business having complete knowledge of the other's.

B. OTHER LICENSORS MAY NOT HAVE ANY KNOWLEDGE OF BROADCASTS

Other businesses license content from content creators (or other licensors) for specific fields of use, and may be completely divorced from broadcast licensing. For example, in some cases, a company acquires sub-licensable rights to content for video-on-demand only (and not for broadcast) and sells those rights to video-on-demand operators. In this case, the licensor (VPO) likely has no knowledge whatsoever of if, when or whether the content they are selling is broadcast over-the-air.

C. VIDEO PROGRAMMING DISTRIBUTORS AND VIDEO PROGRAMMING PROVIDERS SHOULD HAVE "SAFE HARBORS" FOR RELYING ON CERTIFICATIONS OR 3RD PARTY DATA

Under the proposed rules, VPOs must provide either captions or a certification that captions aren't necessary, and the VPD and video programming providers ("VPPs") must transmit the

captions (if supplied) and retain the certifications (if supplied).[7] The Commission also seeks comment on whether and how commercial databases may be used to verify whether captions are required.[8] The rules should also include "safe harbors" for VPD/VPPs that rely upon VPO caption certifications or that rely upon third party data sources.

As discussed above, it may be very difficult for VPOs to know when captioning is required. It is even more difficult for VPD/VPPs to know, as they are unlikely to have any knowledge of other VPD/VPPs' licenses or broadcast rights or broadcast schedule (absent a 3rd party database). As the Commission notes, there are at least two commercial providers of program schedule data which may be helpful in determining whether IP-delivered content must be captioned.[9]

The Commission should establish explicit "safe harbors" for VPD/VPPs, such that a VPD/VPP is not subject to an enforcement action if it relies on a VPO certification in good faith, without actual knowledge that the certification was erroneous.

Alternately (or additionally), the Commission should establish similar safe harbors whereby a VPD/VPP not be subject to an enforcement action if it relies upon comprehensive, commercially-available databases to ascertain whether programming was broadcast with captions after the effective date of the rules. We anticipate that there are few technical or administrative difficulties in making such data available, and such databases would enable automated lookups. If the Commission elects to establish safe harbors or otherwise utilize third party databases, Rovi would be pleased to discuss the details of how such third party databases could operate.

The Commission should not hold VPD/VPPs responsible for a VPO's incorrect certifications, and should not hold VPD/VPPs responsible if they undertake a good-faith effort to determine whether captions are necessary via use of an appropriate third-party database.

III. EFFECTIVE DATES SHOULD TAKE INTO CONSIDERATION WORKFLOW AND RE-PROCESSING EXISTING CONTENT

The Commission proposes adopting the VPAAC recommended schedule of deadlines of six, 12 or 18 months after publication of the rules in the Federal Register, and seeks comment on this

[7] *NPRM* at App. A, proposed § 79.4(c).
[8] *NPRM* at 38.
[9] *NPRM* at n.123 (noting that Rovi and Tribune Media Systems are vendors of comprehensive program schedule data).

schedule.[10] The proposed implementation schedule should be adjusted in light of the considerable undertaking in both processing existing content and developing processes and procedures to ingest new content.

Rovi proposes that each of these deadlines be not less than 18 months after publication of the rules. For all classes of programming (prerecorded and unedited, live or near-live, and prerecorded and edited), content must be ingested into the VPD/VPP's systems and encoded, transcoded, and otherwise processed for a multitude of playback environments and networks. Software and operational processes must be modified to ingest into each of these environments in an automated fashion. Developing and thoroughly testing these systems is a significant undertaking which requires at least 18 months.[11]

Additionally, there is a huge catalog of content currently available via on-demand and similar services which has already been delivered to the VPD/VPPs and ingested into the VPD/VPP's systems.

RES has accumulated content over many years, but which may be broadcast with captions after the effective date and therefore would be obligated under the proposed rules to have captions carried via IP delivery. This obligation would seem to be in effect even if there has been no contact between RES and the VPO for years, and for which RES has neither captions nor a certification that captions are not required.

The Commission should adopt rules or waivers which exempt content which has already been received by a VPD/VPP without captions if no captions have been received from the VPO, and the VPD/VPP has made a good faith effort to notify that VPO that the VPD/VPP does not have caption data and does not have a certification that captions are not required, whether a certification is received from the VPO or not.

The implementation schedule should be set to 18 months after publication in the Federal Register, and should "grandfather" content which has already been ingested without captions.

---

[10] *NPRM* at 28.

[11] Eighteen months is the shortest period, given that certain other of our recommendations are taken favorably. Notably, if the Commission declines to specify SMPTE-TT for caption interchange, Rovi would need to develop and test ingest systems for each interchange system we expect to encounter, which would multiply the amount of development work necessary to prepare and test these systems.

IV. TECHNICAL STANDARDS

A. INTERCHANGE FORMAT

The VPAAC Report suggested "that there be a single standard interchange format for content providers to encode closed captions into programming before they distribute it," such that video programming would not need to be re-captioned to comply with different standards.[12] The Commission seeks comment "on whether to specify a particular standard for the interchange format or delivery format of [closed captions in] IP-delivered video programming."[13]

1. CAPTION INTERCHANGE FORMATS VARY

There are many caption interchange formats. As IP-delivered video programming has proliferated, each programming creator, programming distributor and receiver/decoder manufacturer has made separate and distinct choices as to how (and to what extent) closed captioning is utilized and supported by their respective systems. In this environment, lacking any regulations regarding closed captioning over IP networks, these systems have evolved in varying directions without widely adopted interchange formats – indeed, when the "interchange format involves negotiations between the VPO and the VPD/VPP"[14] there are many factors involved in the negotiation, and facilitating interoperation is not necessarily a primary consideration.

SMPTE-TT (and the broader Timed Text Markup Language (TTML)) was developed specifically to address the incompatibilities between IP video delivery systems, and to avoid requirements for VPOs to generate caption data files several ways for each of several different IP video delivery systems.[15]

2. SMPTE-TT SHOULD BE THE REQUIRED INTERCHANGE FORMAT FOR IP-DELIVERED CONTENT

In order to address the diverse set of interchange formats, SMPTE-TT should be the required interchange format for IP-delivered content. VPOs author and generate closed captioning data in various ways, in various formats. For broadcast content, these caption data are embedded into a video signal using either CEA-608 (for analog) or CEA-708 (for digital) formats. Both

---

[12] *See* VPAAC Report at 17, NPRM at 39. *See also* NPRM at n. 57.
[13] *NPRM* at 40.
[14] *Id.*
[15] The multiple incompatible formats for captions add a significant cost to captioning Internet content. *See* Wikipedia, *Timed Text*, https://secure.wikimedia.org/wikipedia/en/wiki/Timed_Text (as of Oct. 13, 2011, 16:09 GMT).

analog and digital video interchange mechanisms and standards are well-understood by the industry and clearly necessary for one party to supply content to another – that is, for analog and digital video for broadcast and MPVD use, interchange is a well-settled matter that includes an effectively mandated interchange format.

For caption interchange that has not been long subject to a *de facto* interchange format requirement, it is both clear that SMPTE-TT is the appropriate format and that absent a regulation requiring its use for interchange, commercial effects will increase the costs and delay captioning of IP-delivered content.[16]

The Commission proposes to require that captions be provided by VPOs to VPD/VPPs, and require that VPD/VPPs enable the rendering or pass-through of captions.[17] With these requirements, VPD/VPPs are required to accept captions in whatever format a VPO may provide them, and ingest them into its systems – and VPOs have no requirement to supply captions in any particular method. This would create a dynamic where VPD/VPPs are at risk of being forced to pay VPOs for providing useful caption files.

For these reasons, the Commission should adopt SMPTE-TT as the caption interchange format for IP-delivered content.

3. ADOPTING SMPTE-TT AS THE INTERCHANGE FORMAT WILL ACCELERATE DEPLOYMENT OF CLOSED CAPTIONED IP-DELIVERED CONTENT

Adopting SMPTE-TT as the interchange format will minimize the obstacles and development time necessary for proliferation of captioned IP-delivered content. Commercial considerations aside, the Commission should adopt SMPTE-TT as the single interchange format for captions in order to accelerate wide deployment of captioned Internet video. A typical VPD/VPP receives content from various VPOs, and has processes and procedures (both automated and otherwise) that are followed to receive the content, prepare it for distribution and distribute to consumers.

---

[16] However, any caption interchange requirements should be limited to use of SMPTE-TT for carriage of captions utilizing CEA-608 and CEA-708 features. *See infra* at V.B.

[17] *NPRM* at App. A, p. 36 (proposed 47 C.F.R. 79.4(c))

The Commission should adopt SMPTE-TT as the closed caption ingest format, as this would minimize the number of different closed caption ingest processes and allow VPD/VPPs to maximally automate the provision of captions for IP-delivered content.

4. AT THE VERY LEAST, THE COMMISSION SHOULD ADOPT A SMALL SET OF "SAFE HARBOR" INTERCHANGE FORMATS

Alternately, if the Commission declines to specify a single standard for caption interchange, it should at the very least specify a small set of allowable standards for caption interchange. Just as specifying a single standard would minimize the costs, development time and burden, specifying a limited set of standards would put an upper bound on how many different standards VPD/VPPs would need to support – thereby limiting costs, development time and burden.

B. DELIVERY FORMATS

The VPAAC concluded that there should not be a single delivery format, and the Commission has properly declined to specify delivery formats.[18] We support this restraint and believe that declining to require specific delivery formats would enable the largest population of receivers to render captions to consumers. If the Commission were to require specific delivery formats, this would limit the capabilities of the delivery format to a lowest common feature set – and could eliminate the possibility of more advanced features for more advanced devices.

V. APPARATUS

A. ALL APPARATUS

Section 203(a) of the CVAA requires that certain "apparatus designed to receive or play back video programming *transmitted* simultaneously with sound" have caption decoding and display capability.[19] The Commission has sought comment on "what constitutes an 'apparatus'".[20] Covered apparatus should be limited to devices with an implemented, functional primary utility of receiving and playing back *transmitted* audiovisual programming.

The Commission proposes to operate under the assumption that an apparatus designed to receive or play back video is "all hardware *that is used in* receiving or playing back video programming," seeks comment on "how to determine whether hardware is primarily designed for

[18] *See* VPAAC Report at 17, *NPRM* at 39.
[19] Pub. L. No. 11-260, § 203(a) (emphasis supplied).
[20] *NPRM* at 49.

receiving or playing back video programming," whether an apparatus includes software, and how waivers requests should be evaluated and how waivers should be implemented.[21]

1. DEVICES DESIGNED TO RECEIVE OR PLAY BACK VIDEO ARE BECOMING GENERAL PURPOSE COMPUTING PLATFORMS

Video playback and reception products are increasingly being built upon general-purpose computing platforms. Traditionally, video devices have been electronics specifically designed to display video. More recently, devices which had previously been purpose-built to perform this function are becoming much more like general-purpose computing platforms and are enabling additional (Internet-enhanced) functionality described as "Connected TV".[22] Furthermore, devices without purpose-built hardware to receive or play back video programming are increasingly functioning as video receivers. Taken together, we believe that there will be little difference between hardware that is designed for receiving or playing back video programming and general-purpose computing hardware (like desktop PCs) – and this is especially true for devices that receive video programming over the Internet. Over time, it will be difficult or impossible to determine from examining the hardware *only* if it is primarily designed for receiving or playing back video programming.

Therefore, we believe that determining what constitutes an apparatus under Section 203(a) should not be constrained to a determination of the primary purpose of the *hardware design*. It is important to consider the product's actual functionality.

2. DETERMINATION OF APPARATUS SHOULD BE PRIMARILY A SOFTWARE CONSIDERATION

Video reception and playback devices are a combination of hardware and software, and software must enable a feature in order for the product to have that feature. Increasingly, there is little difference between hardware for television-related and hardware for non-television-related functions – especially for IP-connected devices. The amount of processing power and display resolution of general-purpose hardware is often sufficient for television-related applications, and special-purpose television-related hardware is often general-purpose computing hardware.

---

[21] *NPRM* at 49-51.

[22] *See* Smart TV, https://secure.wikimedia.org/wikipedia/en/w/index.php?title=Smart_TV&oldid=454447222 (last visited Oct. 14, 2011).

Therefore, in order to determine if a device is an apparatus under Section 203(a), the Commission should consider whether the *product sold* is designed to receive or play back video programming transmitted simultaneously with sound, considering both hardware and software, by considering at least several factors:

First, whether the hardware is capable of receiving "video programming transmitted simultaneously with sound", decoding and rendering the video and sound received, and producing the audio and sound on an output or a display.

Second, whether the hardware and software, taken together, actually enables the reception of video programming transmitted simultaneously with sound, decoding and rendering the video and sound received, and producing the audio and sound on an output or a display.

Third, whether the product is designed for this feature or whether the capability is ancillary to the product's design as evinced by, *e.g.*, the manner the product is sold and marketed, the product's primary functionality, or whether the product's "essential utility" is related to audio/video reception and display.

Importantly, each of these factors should be included in the Commission's evaluation of a product and definition of apparatus, which will illuminate: when products are "primarily designed for activities other than receiving an playing back video programming transmitted simultaneously with sound;"[23] those products with "essential utility [] derived from other purposes;"[24] and if a product is "designed to receive or play back video programming transmitted simultaneously with sound."[25]

B. REGULATION SHOULD ESTABLISH MINIMUM FUNCTIONALITY WHILE ENABLING COMPETITION AND INNOVATION IN ACCESSIBILITY FEATURES

Regulations implementing Section 203(a) should establish a minimum functionality for devices (if achievable), but should not limit innovation or additional features. Rovi's products support rendering of closed captioning on multiple devices, such as TVs, phones, and PCs and

[23] Pub. L. No. 11-260, § 203(a).
[24] *Id.*
[25] *Id.*

believes that generally such requirements are not onerous for the playback of video programming delivered via the Internet. The VPAAC Report recommends mirroring the feature set of existing television receivers,[26] and the Commission proposes rules to that effect.[27]

The Commission should clarify that this functionality is a minimum set, and that devices may exceed these required features and innovate in accessibility features. But this set must be a minimum, as many devices have processing or power limitations that make implementation of complex rendering mechanisms infeasible and unachievable. Only baseline functionality should be generally mandated, if at all.

In particular, should the Commission adopt SMPTE-TT as a required interchange format (as we recommend *supra*), the functionality of SMPTE-TT which is required for compliance should be limited to encapsulation and carriage of CEA-608 and CEA-708 caption functionality.[28]

VI. CONCLUSION

The VPO may not be the most likely to know whether captions are required, and the owner of the video programming may not have a direct business relationship with the VPD/VPP, therefore the Commission should have a "safe harbor" for relying upon either certifications and/or third party data.

In order to minimize costs, burdens and implementation time, the Commission should adopt SMPTE-TT as the interchange format for captions, limited to CEA-608 and CEA-708 functionality. In the event the Commission concludes that specifying a single interchange format, it should specify a small set of interchange formats, to limit costs, burdens and implementation time. The Commission should not specify caption delivery formats for IP-delivered programming.

Apparatus covered by the regulation should be limited to devices that receive or play back programming which is transmitted. Furthermore, such apparatus should be limited to devices which are capable of receiving or playing back IP-delivered programming, and have such capability enabled, and such capability is not ancillary to the device's essential utility.

---

[26] VPAAC Report at 13-16.
[27] *NPRM* at 56.
[28] SMPTE-TT is a much broader standard which contains many other features and functionality for other (non-television) use, which should not be required by regulation.

Finally, the Commission should "grandfather" certain programming which have been received and ingested by VPD/VPPs long before the effective date, but which were never supplied with caption data.

Respectfully submitted,

Adam Powers
Vice President of Technology
Rovi Corporation
2830 De La Cruz Blvd
Santa Clara, CA 95050
apowers@rovicorp.com

October 18, 2011